



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) EP 0 806 733 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
12.11.1997 Bulletin 1997/46

(51) Int. Cl.⁶: G06F 17/30

(21) Application number: 97107090.9

(22) Date of filing: 29.04.1997

(84) Designated Contracting States:
DE FR GB

(72) Inventor: Stone, Harold S.
Princeton, NJ 08540 (US)

(30) Priority: 06.05.1996 US 643691

(74) Representative:
VOSSIUS & PARTNER
Siebertstrasse 4
81675 München (DE)

(71) Applicant: NEC CORPORATION
Tokyo (JP)

(54) Method and apparatus for multi-resolution image searching

(57) A multiresolution method and apparatus for searching of a database of images where the search is performed on compressed images, without first decompressing them. The method searches the database of compressed images first at a low resolution to obtain the relative quality of a match between a search template and a candidate image. If the match is below a particular threshold, the search is terminated without committing any further computational resources to the

search. Conversely, if the match is above a particular threshold, the method enhances the resolution of the candidate image and then performs another match. As long as the relative quality of the match is above the particular threshold, the resolution of the candidate image is successively enhanced, until a match determination is made at a full resolution of the candidate image.

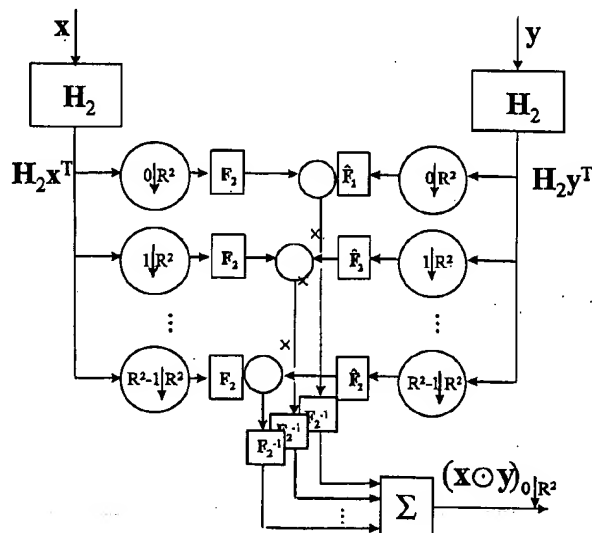


Figure 4

EP 0 806 733 A2

Description

This invention relates to the field of image storage and retrieval and in particular to a method and apparatus for searching for a particular image contained within a database of images.

Image databases are databases that contain digital data of images and are finding ever-increasing uses in business and entertainment. The widespread use of image databases, however, is intensifying the need to develop more effective ways of searching such databases.

By way of background, a database is a collection of related data. Oftentimes, the data in a database are structurally organized as tables, fields and records. Generally, each record in the database possesses a set of attributes and a user may wish to conduct a search of the database based upon a value of one or more of the attributes. For example, in a database of football statistics, one may desire to search for all quarterbacks who completed 100 or more passes in a single season, etc.

In conducting such a search, it is assumed that each record possesses several attributes and the user wants to search for records that possess certain values of certain attributes. The specification of which records are desired is called a query and is usually restricted to one of three types: a) A simple query which specifies a specific value for a specific attribute; e.g., THROWINGARM = LEFT; or THROWINGARM = RIGHT; b) a range which specifies a specific range of values for a specific attribute; e.g., WEIGHT < 220 or 180 < WEIGHT < 220; and c) a BOOLEAN such as (THROWINGARM = LEFT) AND (HEIGHT > 6') AND (COMPLETIONS > 99).

Unfortunately, such query-based methods do not solve the problem of searching an image database because image data is oftentimes not annotated with attribute descriptions. Consequently, an increasing interest and continuing need exists for developing search and retrieval methods which efficiently, quickly and accurately searches a database of images.

Consequently, the prior art has shown an increasing interest in searching image databases by image content. See, for example, M. Flickner et al., "Query by Image and Video Content: The QBIC System", Computer, Vol. 28, No. 9., pp. 23-32, September 1995. To perform such a content-based image search, one would typically select an image template and query the database to find images that "look like" the image template. An alternative form of this question is "where in a waveform B does a signal A occur?"

In providing an answer to this latter question, the prior art has utilized correlation techniques from signal-processing to measure similarities of images in much the same way that they are used to measure the similarities of signals. Such correlation techniques require the computation of the circular correlation of an image x with a search pattern y where the circular correlation is defined to be the mathematical operation given by the formula:

$$(x \cdot y)_j = \sum_{i=0}^{i=N-1} x_{(i+j) \bmod N} y_i.$$

As used throughout the present application, input variables are denoted by lower case letters x and y where x is one of several possible images and y is a fixed pattern or subimage sought in images contained in an image database. Note that while images are two dimensional, all of the discussion herein is for one-dimensional row vectors x and y thereby simplifying the notation used. Those skilled in the art will readily appreciate that the extension to two dimensions is straightforward. Additionally, and unless explicitly stated otherwise, the x and y vectors both have a length depicted as N . For search operations, if the sought pattern y is shorter than x , then, y is padded to length N by adding 0s to its length. Finally, x^T is used to denote the transpose of x , which changes the row $N \times 1$ row-vector x into a $1 \times N$ column vector.

The fundamental operation for search operations is the circular correlation of x to y which is defined to be the mathematical operation given by the formula:

$$(x \cdot y)_j = \sum_{i=0}^{i=N-1} x_{(i+j) \bmod N} y_i.$$

Typically, the circular correlation is used together with other functions such that a new function is produced whose peaks indicate positions within x that best match y . In operation, the circular correlation produces a vector of length N from two vectors of length N .

As used herein, the notation $(x \cdot y)_{i \downarrow R}$ denotes a subsampling of the correlation operation by taking components whose indices are equal to i modulo R . For example, if $R = 4$ and $i = 1$, the subsampling operation extracts elements 1, 5, 9, 13, ... from the circular correlation of x and y . Consequently, $(x \cdot y)_{i \downarrow R}$ is called subband i of $(x \cdot y)$.

The invention of the present application advantageously uses wavelet transforms of images. A wavelet transform of an image x is for a wavelet transform H , where H is an $N \times N$ matrix, denoted by a N -vector $H x^T$. The wavelet transform is advantageous for image compression because it leads to transformed vectors with many components of small magnitude, which can be represented by a relatively small number of bits. Often these components are replaced by 0s which are efficiently encoded in the wavelet transform. Those skilled in the art can readily appreciate that not only are the image representations reduced, but the time required to process the wavelet transforms is reduced as well because any zero components do not have to be processed during arithmetic operations.

In the discussion that follows, it is assumed that the wavelet transform matrix H has a special structure. Specifically, it is assumed that H is *block diagonal* with a block size of R . By definition, this means that H may be partitioned into individual blocks having a size $R \times R$, and that all of the blocks within H — except for the individual blocks that fall on the diagonal contain only 0s. By way of example, the matrix Y below is a block diagonal matrix having a block size of 3. Its diagonal blocks are identical to each other, and the following discussion is limited to such block diagonal matrices which have identical blocks down their diagonal.

$$Y = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 & 1 \end{bmatrix}$$

In an article entitled "Orthonormal and Biorthonormal Filter Banks as Convolvers and Convolutional Coding Gain", which appeared in IEEE Trans. On Signal Processing, Vol. 41, No. 6., pp. 2110-2130, June 1993, P. Vaidyanathan proved a convolutional theorem for transformed vectors (hereinafter referred to as "Vaidyanathan's theorem"). Vaidyanathan's theorem states that given wavelet transforms of x and y , then some of the components of the circular convolution of x and y may be found without first transforming x and y from the wavelet domain back to the pixel domain. Graphically, the theorem may be depicted as in Figure 1.

According to this theorem and with reference to Figure 1, when H is block diagonal having a block size R with identical blocks along the diagonal,

$$H^T H = I_N,$$

and where I_N is the identity matrix of size N , then subband 0 of the circular correlation of x and y is obtained from the wavelet transforms of x and y , by computing the circular correlations of all subbands of x and y and adding these correlations together, component by component.

Note that the computation does not require that the wavelet transforms be inverted. For example, if $N = 32$ and $R = 4$, each subband contains 8 components, and there are four such subbands. Vaidyanathan's theorem requires the formation of 4 subbands of x and y , the calculation of 4 circular correlations each of length 8, and then the addition of these together component by component to produce a single vector of length 8. This resulting single vector is exactly equal to subband 0 of the circular correlation of the full vectors x and y .

Subband 0 of the circular correlation gives only a low-resolution version of the circular correlation. And while it may indicate that a match for a particular search may be present, there remains the possibility that no match exists. Therefore, an effective search method must be able to produce the remaining components of the circular correlation to resolve if a match is present. Conveniently, Vaidyanathan's theorem may be extended to the computation of other subbands of a circular correlation as indicated in Figure 2.

With further reference to Figure 2, to compute Subband j , the vector x is cyclically shifted to left by j positions, and then processed by the same mechanism that computes Subband 0. One overwhelming problem with this approach, however, is that it is excessively computationally expensive. For example, if it is assumed that the "cost" of computing the wavelet transform and the inverse wavelet transform are about equal, then the prior art method depicted in Figure

2 indicates that the cost of computing all of the subbands of the cyclic correlation of x and y is the cost to compute R direct wavelet transforms for x and one for y , together with the cost of the cyclic correlations.

However, the cyclic correlation of x and y may be computed by an alternative method, namely by taking inverse transforms of Hx^T and Hy^T , followed by cyclic correlations in the pixel domain. This alternative method incurs a cost equivalent to two wavelet transforms plus the cost of the cyclic correlations. Vaidyanathan's method requires $R-1$ more wavelet transforms. The net cost comparison depends on the relative cost of the cyclic correlations when done by Vaidyanathan's method and when done in the pixel domain by inverting the transforms first.

Clearly, the cost of the cyclic correlations depends on how they are computed. As indicated by Figure 3(a), the well-known convolution theorem provides an efficient way to compute these correlations. With reference to Figure 3(a), there it is shown a method for computing the circular correlation of x and y . A straightforward implementation of this computation requires N^2 multiplications. A more efficient implementation for sufficiently large values of N is shown in Figure 3(b). As shown in Figure 3(b), to compute the circular correlation of x and y , one must first, compute their Fourier transforms; second, multiply the Fourier coefficients X_i and Y_i , point by point in the frequency domain; and third, take the inverse Fourier Transform of the point-by-point product. (Note that the notation \hat{F} shown in Figure 3(b) denotes taking the complex conjugate of a Fourier Transform, and incurs a cost equal to that of taking the Fourier transform itself).

As one skilled in the art can readily appreciate, the Convolution Theorem improves the efficiency of the calculation because each forward or inverse transform requires only about $5 N \log N$ operations, and the point-by-point multiply requires only N . This is usually much less than the N^2 operations required to compute the circular convolution by more naive methods. To compare Vaidyanathan's method to the straightforward one, the costs of each method are computed as follows:

Straightforward Method:

- 1) Invert Hx^T and Hy^T . (Cost: 2 wavelet transforms).
- 2) Calculate Fourier Transforms of x and y (Cost: $2 N \log N$ operations).
- 3) Do the point-by-point multiplies. (Cost: N operations).
- 4) Transform back to the pixel domain. (Cost: $N \log N$ operations).

Total cost: 2 wavelet transforms plus $3 N \log N + N$ operations.

Vaidyanathan's method:

- 1) For each subband, calculate the wavelet transform of x . (Cost: R wavelet transforms, or $R-1$ wavelet transforms if Subband 0 is available already).
- 2) For each subband, calculate R cyclic correlations of length N/R . (Cost: per subband it is $N \log N - N \log R$ operations; in total it is $RN \log N - RN \log R$).

As can be readily seen, the total cost for Vaidyanathan's method is about R times as large as the straightforward method, which moves first to the pixel domain and then does fast correlation there. The straightforward method is not progressive. It requires a full calculation to obtain any significant search result. And while Vaidyanathan's method is progressive and can be discontinued before performing all of the operations, it is simply too computationally expensive to use in practice.

Consequently, a continuing need exists in the art for methods and apparatus which conveniently, efficiently and inexpensively permit the searching of a large number of candidate images contained within an image database.

The above problems are solved and an advance is made over the art according to the principles of the present invention whereby images are highly-compressed through the use of wavelet compression techniques, and Fourier correlation techniques are utilized to perform a high-speed search and retrieval of images contained in an image database.

The method advantageously permits the searching of a database of images where the search is performed on the compressed images, without first decompressing them. Specifically, the method of the present invention operates by searching the database of compressed images first at a low resolution to obtain the relative quality of a match between a search template and a candidate image. If the match is below a particular threshold, the search may be terminated without committing any further computational resources to the search. Conversely, if the match is above a particular threshold, the method enhances the resolution of the candidate image and then performs another match. As long as the relative quality of the match is above the particular threshold, the resolution of the candidate image is successively enhanced, until a match determination is made at a full resolution of the candidate image.

Viewed from another aspect, the present invention provides an apparatus for the efficient searching of images contained in an image database. Further features and advantages of the present invention, as well as the structure and operation of various embodiments of the present invention are described in detail below with reference to the accompanying drawings.

The teachings of the present invention can be readily understood by considering the following detailed description in conjunction with the accompanying drawings, in which:

- Figure 1 is an illustration of a prior art method for generating subband 0 of a circular correlation;
- Figure 2 is an illustration of a prior art method for generating other subbands of a circular correlation;
- Figure 3 is an illustration depicting a convolution method;
- Figure 4 is an illustration of the first step of a multiresolution method in accordance with the present invention;
- Figure 5 is an illustration of the multiresolution method performed in a Fourier-wavelet domain in accordance with the present invention;
- Figure 6 is an illustration of a multiresolution method performed on right block-circulant matrices; and;
- Figure 7 is a flow diagram showing the steps performed for a multiresolution image search according to the present invention.

A preferred embodiment of the invention will now be described while referring to the figures, several of which may be simultaneously referred to during the course of the following description.

The invention of the present application utilizes a multiresolution technique which overcomes the debilitating inefficiency associated with the prior art. By way of additional background, in multiresolution wavelet decompositions there exist a sequence of matrices H_1, H_2 , etc. such that each successive matrix produces finer and finer wavelet transforms. In the following example, H_2 is a coarse transform that is block diagonal with blocks of size R^2 , and operates on R^2 subbands, each of length N/R^2 . H_1 is a fine transform matrix that is block diagonal with blocks of size N/R , and operates on R subbands, each of length N/R . There exists an update matrix U_{12} , that refines H_2 into H_1 by the equation:

$$H_1 = U_{12} H_2.$$

The matrix U_{12} is block diagonal with blocks of size R^2 . The blocks of U_{12} consist of R interlaced copies of the block matrix that is on the diagonal of H_1 . For example, let H_2 be the coarse wavelet transform given by:

$$H_2 = 0.5 \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \end{bmatrix}$$

The fine transform is given by the matrix H_1 whose structure is given by:

$$H_1 = \sqrt{2}/2 \cdot \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

The update matrix U_{12} is the matrix given by:

$$U_{12} = \sqrt{2}/2 \cdot \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix}$$

Note that the pattern of 1s and -1s in the one-step update matrix U_{12} is the same as in the fine wavelet transform matrix except that the 1s and -1s in U_{12} are interlaced every 2 apart and they lie within blocks of size 2 in H_1 and within blocks of size 4 within U_{12} .

With reference now to Figure 4, there it shows the structure of the computation when the convolution theorem is applied to Vaidyanathan's theorem for a wavelet transform H_1 that is block-diagonal with identical blocks of size R on the diagonal.

Specifically, the R circular correlations are replaced by Fourier transforms F_2 that operate on vectors of size N/R . The component-by-component multiplies are followed by inverse Fourier transforms to produce Subband 0 of the circular correlation. Other subbands may be computed by applying this process recursively in a multiresolution analysis. The process is efficient because the cost in computation is no greater than the cost of inverting the wavelet transforms first, and following with the efficient implement of circular correlations in the pixel domain by means of Fourier domains in that domain and component-by-component multiplication.

For a multiresolution analysis, the coefficients of Subband 0 mod R^2 are obtained as indicated in Figure 4. There are N/R^2 such coefficients in each of R^2 subbands. Subsequently, the relation represented by Figure. 5 is applied thereby obtaining the coefficients of Subband 0 mod R . There are N/R of these coefficients, or R times as many as in the coarse analysis.

With further reference to Figure 5, there it shows two equivalent operations. In the upper part of Figure 5 it shows the subsampling of the H transform followed by a coarse Fourier transform (F_2), and then shows how to obtain the Fourier transforms for finer resolutions from the coarser ones. As shown in the Figure, the R^2 vectors of length N/R^2 are transformed by the matrices u_{12} . Each of these matrices is a square matrix of size N/R^2 and is a block of the update

matrix U_{12} that refines H_2 into H_1 .

Figure 5 shows that the coarse transforms are upsampled. The upsampling interlaces together R^2 vectors of length N/R^2 to create R vectors of length N/R . (Interlacing the vectors (1,3,5,7), (2,4,6,8), (9,11,13,15), by (10,12,14,16) by upsampling by 2, produces the vectors (1,9,3,11,5,13,7,15) and (2,10,4,12,6,14,8,16).)

After the upsampling, and as shown in the Figure, the resulting vectors are transformed by the matrix T_{12} . This matrix updates a fourier transform that operates on interlaced vectors of length N/R^2 to a fourier transform that operates on vectors of length N/R . That is, T_{12} satisfies the equation:

$$F_1 = T_{12}(I_R \times F_2).$$

The "cross" operation in this equation indicates a matrix cross product, and in this case expresses the fact that F_2 is a fourier transform that is repeated and interlaced with a period of length R to create a matrix F_1 of size $N/R \times N/R$ from a matrix F_2 having size $N/R^2 \times N/R^2$. The number of operations required to perform T_{12} on the output of the interlaced transforms F_2 is proportional to N/R . Those skilled in the art will recognize that the computation at the top is exactly equal to the computation at the bottom. That is, the effect of the fine transform H_1 followed by the fourier transform F_1 may be computed by using the update matrices u_{12} and the fourier update matrix T_{12} on the fourier transforms of the coarse wavelet transforms. The amount of work is substantially less by the computation at the top because the Fourier update matrix requires only N/R operations instead of $N/R \log N/R$ operations. Moreover, it is not necessary to compute the full inverse fourier transform in Fig. 4 to obtain the fine subbands of the circular correlation. It is necessary only to compute the subbands that have not yet been computed, and this requires less work. It requires one overall step with a number of operations proportional to N/R to do the global operations that affect all subbands, and then just the operations for the subbands that have not been computed. If $R = 2$, then half of the outputs are in Subband 0, which is already known, and only half of the remaining computation of the inverse fourier transform has to be performed.

Advantageously, the multiresolution method of the present invention is recursive and can be repeated on any outputs shown in Figure 5 to create a full resolution search. If the full resolution search is carried out, the amount of computation performed is precisely the same as if an inverse wavelet transform had been performed on the original image transforms, and then followed by a fourier transform from the pixel domain to the frequency domain, point-by-point multiplication in the frequency domain, and an inverse fourier transform back to the pixel domain.

Figure 7 provides an overview of the steps necessary to perform a multiresolution search according to the present invention. Specifically, for a particular image a pattern template is transformed into a wavelet pattern at block 70. A candidate image is then coarse correlated with the transformed pattern at block 71 and a determination is made at block 72 whether the transformed pattern suitably matches the candidate image. If not, then another candidate image may be chosen or the search abandoned, 79. If the match was suitable, then the candidate image is medium correlated 73 and another similar match test is performed at block 74. The procedure proceeds similarly through blocks 75 and 77 with fine correlated and fully correlated images respectively undergoing a suitable match check at blocks 77 and 78. If a suitable match is found for the fully correlated image at block 78, then the image searched for has been found.

In the discussion below, the following definitions are used:

1. The matrix I_N is an identity matrix of size $N \times N$.
2. The notation $A \times B$ denotes the matrix cross product of matrices A and B , and is sometimes called the Kronecker product.
3. The matrix F_N is a Fourier transform matrix that operates on vectors of length N .
4. The matrix $F_{M,R}$ is an interlaced Fourier transform matrix with the structure $F_M \times I_R$, that is it has R interlaced copies of transforms of size M .
5. The matrix $T_{N,M,R}$ is a Fourier update matrix that transforms $F_{M,R}$ into F_N . That is,

$$F_N = T_{N,M,R} F_{M,R}.$$

Those skilled in the art will recognize that $T_{N,M,R}$ can be implemented by means of one stage of R -way butterfly operations.

6. The matrix C_8 is an 8×8 discrete cosine transform (DCT) matrix that is used to create transforms of 8×8 subimages in a JPEG representation of an image.
7. The matrix H is an $N \times N$ matrix with the structure $I_M \times C_8$, where $N = 8M$. The H matrix produces the JPEG transform of an image vector of length N .
8. The matrix W is a 2×2 Haar transform matrix. It is given by:

$$W = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

6. The matrix

$$V_2 = L_1 \times W$$

consists of 4 interlaced copies of W , and is of size 8×8 .

7. The matrix $V_{4,2}$ has the structure

$$V_{4,2} = I_2 \times (W \times I_2).$$

Those skilled in the art will recognize that the matrix

$$V_4 = I_2 \times (W \times W)$$

satisfies the equality

$$V_4 = V_{4,2} V_2.$$

8. The matrix $V_{8,4}$ satisfies the equality

$$C_8 = V_{8,4} V_{4,2} V_2.$$

Those skilled in the art can confirm that $V_{8,4}$ satisfies the equation

$$V_{8,4} = V(W \times L_1)$$

where V is the matrix below:

$$V = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.180 & 0 & 0.213 & 0 & 0.318 & 0 & 0.906 \\ 0 & 0.375 & 0 & 0.768 & 0 & -0.513 & 0 & -0.075 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.0906 & 0 & -0.318 & 0 & 0.213 & 0 & -0.180 \\ 0 & 0 & 0.383 & 0 & 0 & 0 & 0.924 & 0 \\ 0 & 0 & 0.924 & 0 & 0 & 0 & -0.383 & 0 \\ 0 & -0.075 & 0 & -0.513 & 0 & 0.768 & 0 & -0.375 \end{bmatrix}$$

For a given image stored in JPEG format, a preferred multiresolution analysis that uses the invention is described as follows.

- 1) To correlate an N -vector image x stored as a JPEG transform Hx^T with an instance of a pattern y , extend the pattern to the same length as x and compute its JPEG transform Hy^T .
- 2) Coarse correlation - Generate the Fourier transforms $F_{M,8}Hx^T$ and $\hat{F}_{M,8}Hy^T$. Multiply the transforms point-by-point and partition them into the appropriate number of subbands of length M . Add these vectors, and take the inverse Fourier transform of the sum. For this example, every eighth point of the correlation is generated.
- 3) Medium correlation - Multiply $F_{M,8}Hx^T$ by

$$(T_{2M,M,8} \times I_4) (W \times L) V^{-1}$$

and $\hat{F}_{M,8}Hy^T$ by

$$(\hat{T}_{2M,M,8} \times I_4) (W \times L) V^{-1}$$

- thereby producing four interlaced Fourier transforms of size $2M$. Multiply the resulting vectors point-by-point, and partition them into four subbands. Add the subbands, thereby creating a single vector of length $2M$. Take the inverse Fourier transform of size $2M$, but after a first butterfly, operate only on elements that produce an odd component of the result. This yields a correlation at indices that are multiples of 4 mod 8 of the full correlation.
- 4) Fine correlation - Multiply the x and y transforms of the preceding step by $(T_{4M,2M,4} \times I_2)V_{4,2}$ and $(\hat{T}_{4M,2M,4} \times I_2)V_{4,2}$, respectively. Each vector consists of two interlaced Fourier transforms of size $4M$. Multiply these transforms point-by-point, partition into two subbands and add the subbands. Take the inverse Fourier transform of the result, and after the first butterfly stage, compute only the terms that contribute to the odd results. This produces the components of the correlation whose indices are equal to 2 mod 8 and 6 mod 8.
- 5) Full resolution - For the final step, multiply the x and y transforms from the last step by $T_{8M,4M,2}V_2$ and $\hat{T}_{8M,4M,2}V_2$, respectively. Repeat the analysis noting that there is a single subband of length $N = 8M$ in the transform vectors and in the transform of the correlation product.

At this point, a number of advantages resulting from the multiresolution analysis of the present invention are apparent to those skilled in the art. Specifically, if the search is not going to succeed, and this is discovered early, the search can be terminated early in the multiresolution analysis, because match information is obtained early in the search. If the transforms to the image domain are done first rather than performing multiresolution analysis, then the work is committed before any information is used, and the cost cannot be avoided later as search information is discovered. The wavelet representations tend to be sparse, that is, they contain many zeros, and this can reduce computational cost. The pixel domain representation is not sparse, because virtually all the pixels are nonzero. Consequently, it may be possible to exploit the sparsity in the wavelet domain and equivalent exploitation is not possible in the pixel domain.

Those skilled in the art can appreciate that these results may be readily extended to more general matrices. Specifically, the results extend to *right block-circulant matrices*. A right circulant matrix X is a square matrix in which each row is the same as the immediately preceding row, except that it is shifted cyclically to the right by 1 position. For example, the matrix X below is right circulant.

$$X = \begin{bmatrix} 0 & 5 & 2 & 7 \\ 7 & 0 & 5 & 2 \\ 2 & 7 & 0 & 5 \\ 5 & 2 & 7 & 0 \end{bmatrix}$$

A right block-circulant matrix is a matrix composed of equal-sized submatrices (blocks), and each row of blocks is the same as previous row of blocks, except that it is shifted to the right by the size of 1 block. The matrix Y below is right-block circulant with a block size of 2.

$$Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ -2 & 0 & 0 & 0 & 1 & -1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & -2 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

The equality depicted in Figure 5 holds for block diagonal wavelet transform matrices whose diagonal blocks are equal, and they must satisfy the equation

$$H H^T = H^T H = I.$$

Although these are common wavelet transforms, as, for example, the Haar wavelet, they do not represent the broadest class of wavelets in use. The wavelets of most interest have a transform equation that is right block-circulant, as, for example, the Debauchies wavelets. The modified form of Figure 5 for such wavelets is computationally just as simple, but it is slightly different and is given in Figure 6.

The principle difference between the two figures lies in the update matrices u_{12} and \hat{u}_{12} . In Figure 5, u_{12} is a sub-block of the update matrix U_{12} , and can be obtained directly that matrix. Moreover, each instance of u_{12} in Figure 5 uses the same matrix u_{12} for the computation. In Figure 6, the matrix \hat{u}_{12} has to be calculated from U_{12} , and the entries in the matrix \hat{u}_{12} depend on the subband that it transforms. The matrices \hat{u}_{12} can be computed by the formula:

$$\hat{U}_{12} = (F_2 \times I_{R^2}) U_{12} (F_2^{-1} \times I_{R^2})$$

where F_2 is the Fourier transform in Figures. 5 and 6 that transforms vectors of length N/R^2 and the matrix cross product operation interlaces R^2 copies of this matrix together across R^2 vectors of length N/R^2 . The matrix \hat{U}_{12} is block diagonal, and within its blocks along diagonal lines it contains Fourier transforms of the rows of U_{12} . For a search of one pattern over many images of the same size, the matrices \hat{U}_{12} can be computed once and stored reused when searching each image.

Claims

1. A content-based image search method comprising the steps of:

5 transforming a pattern-template such that the transformed pattern template may be correlated with a wavelet candidate image;
 coarse-correlating said candidate image in the Fourier domain; and
 comparing said coarse-correlated candidate image with said pattern-template to determine if there exists a suitable match.

10

2. The method according to claim 1 further comprising the step of:

 medium-correlating said coarse-correlated candidate image; and
 comparing said medium-correlated candidate image with said pattern-template to determine if there exists a suitable match.

15

3. The method according to claim 1 or 2, further comprising the step of:

 fine-correlating said medium-correlated candidate image; and
 comparing said fine-correlated candidate image with said pattern-template to determine if there exists a suitable match.

20

4. The method according to claim 1, 2 or 3, further comprising the step of:

25 fully-correlating said fine-correlated candidate image; and
 comparing said fully-correlated candidate image with said pattern-template to determine if there exists a suitable match.

25

5. An apparatus for performing a content-based image search comprising:

30

 means for transforming a pattern template into a transformed pattern template such that said transformed pattern template may be correlated with a wavelet candidate image;
 means for coarse-correlating said candidate image in a Fourier domain; and
 means for comparing said coarse-correlated candidate image with said transformed pattern-template to determine if there exists a suitable match between the candidate image and the pattern template.

35

6. An apparatus for performing a content-based image search according to claim 5 further comprising:

 means for medium-correlating said coarse-correlated candidate image; and
 means for comparing said medium-correlated candidate image with said transformed pattern-template to determine if there exists a suitable match between the candidate image and the pattern template.

40

7. An apparatus according to claim 5 or 6, further comprising:

45 means for fine-correlating said medium-correlated candidate image; and
 means for comparing said fine-correlated candidate image with said transformed pattern-template to determine if there exists a suitable match between the candidate image and the pattern template.

45

8. An apparatus according to claim 5, 6 or 7, further comprising :

50

 means for fully-correlating said fine-correlated candidate image; and
 means for comparing said fully-correlated candidate image with said transformed pattern-template to determine if there exists a suitable match between the candidate image and the pattern template.

55

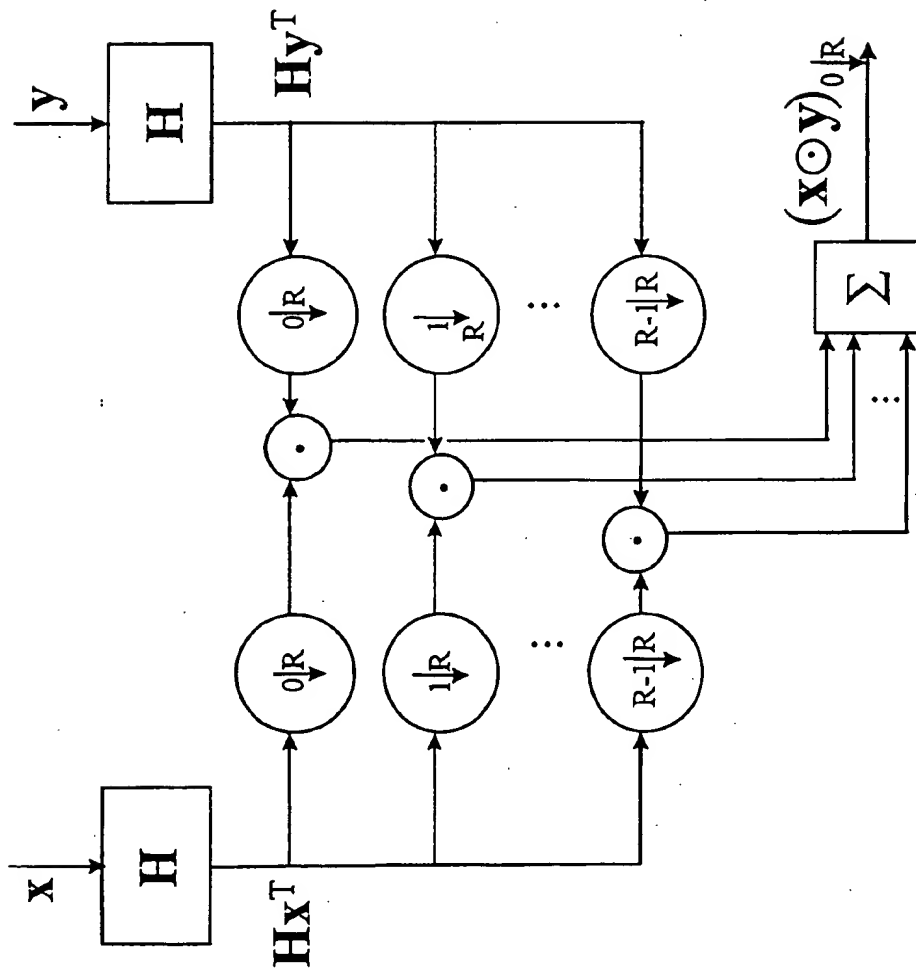


Figure 1
Prior Art

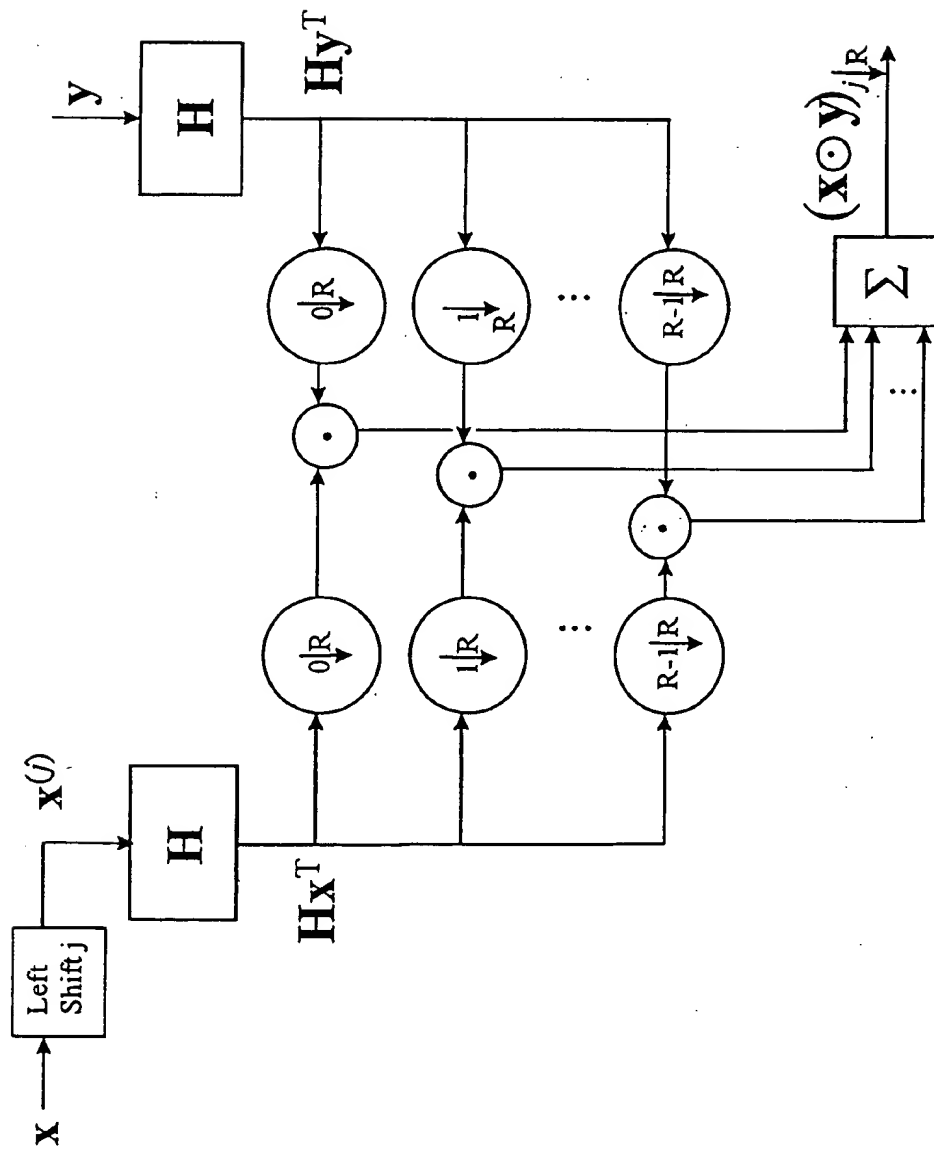


Figure 2
Prior Art

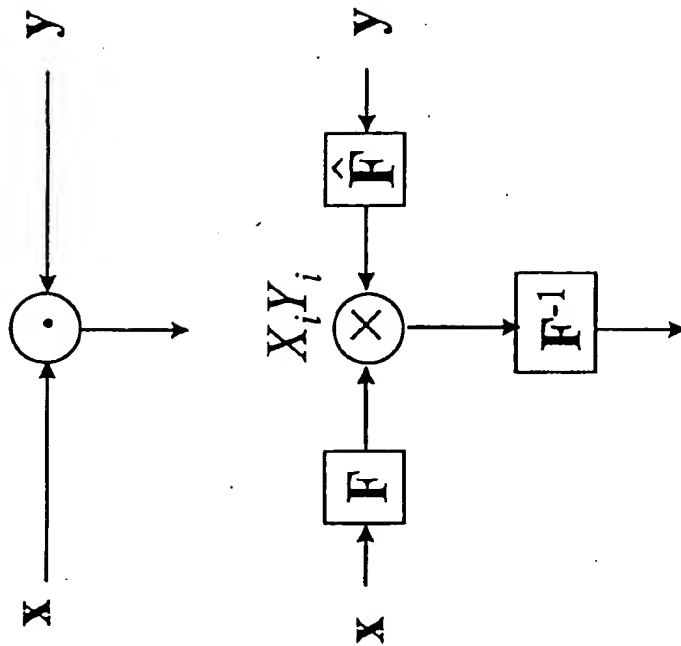


Figure 3
Prior Art

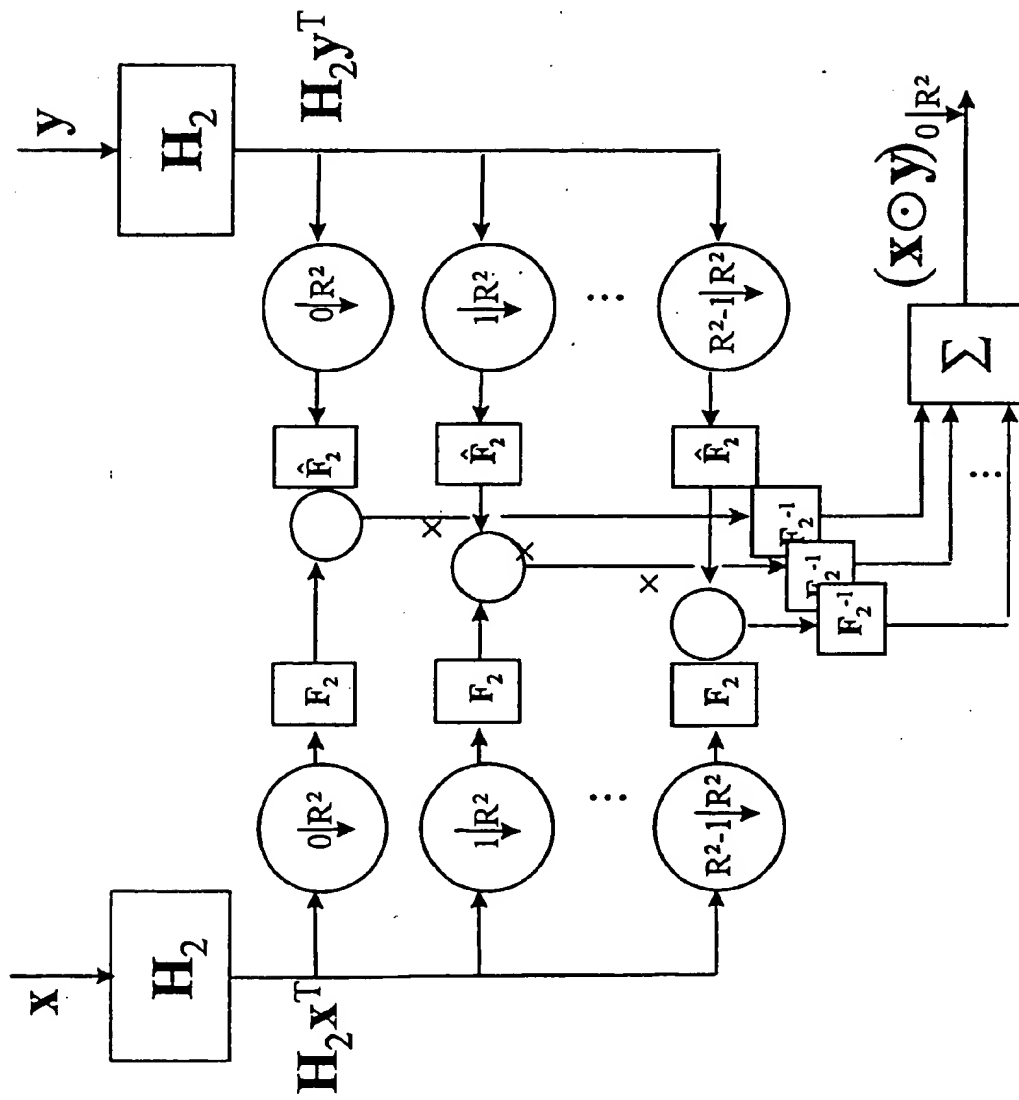


Figure 4

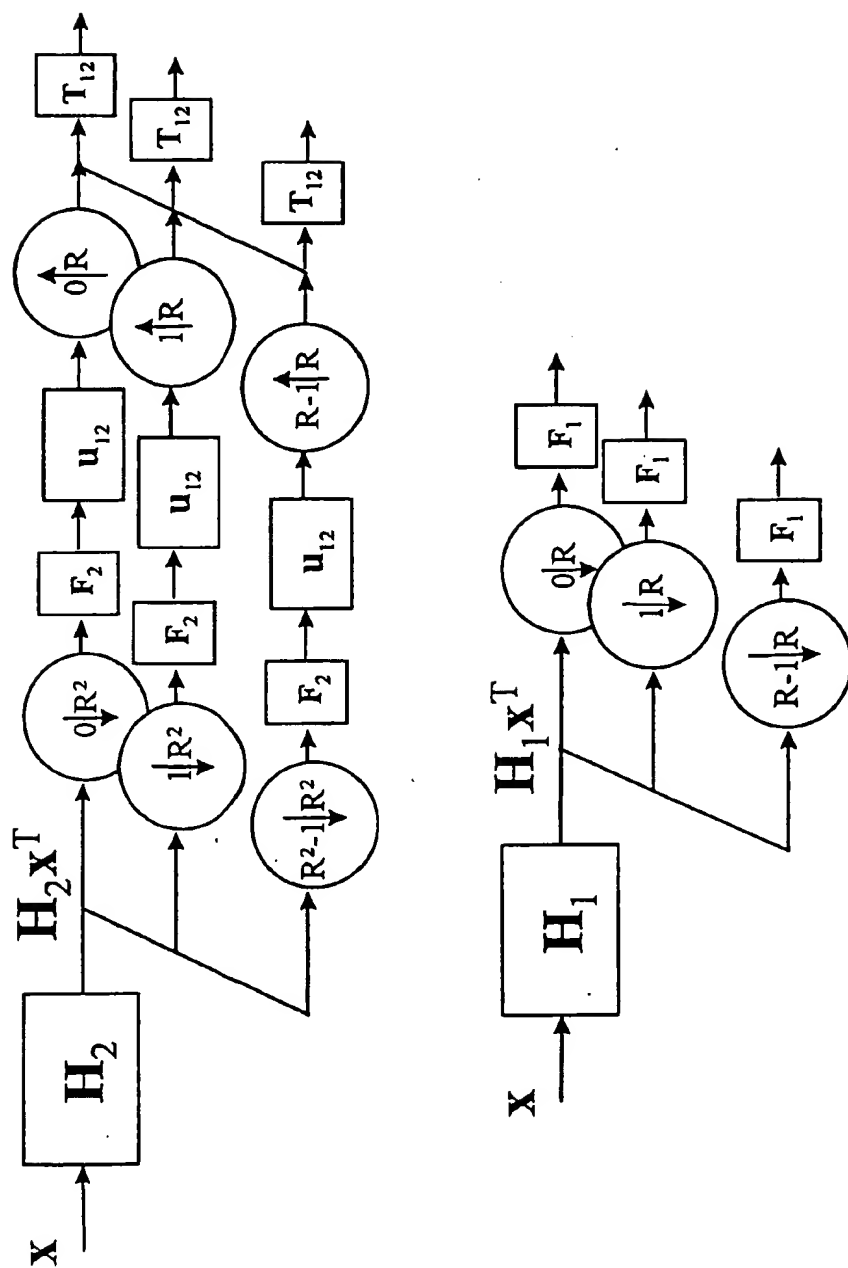


Figure 5

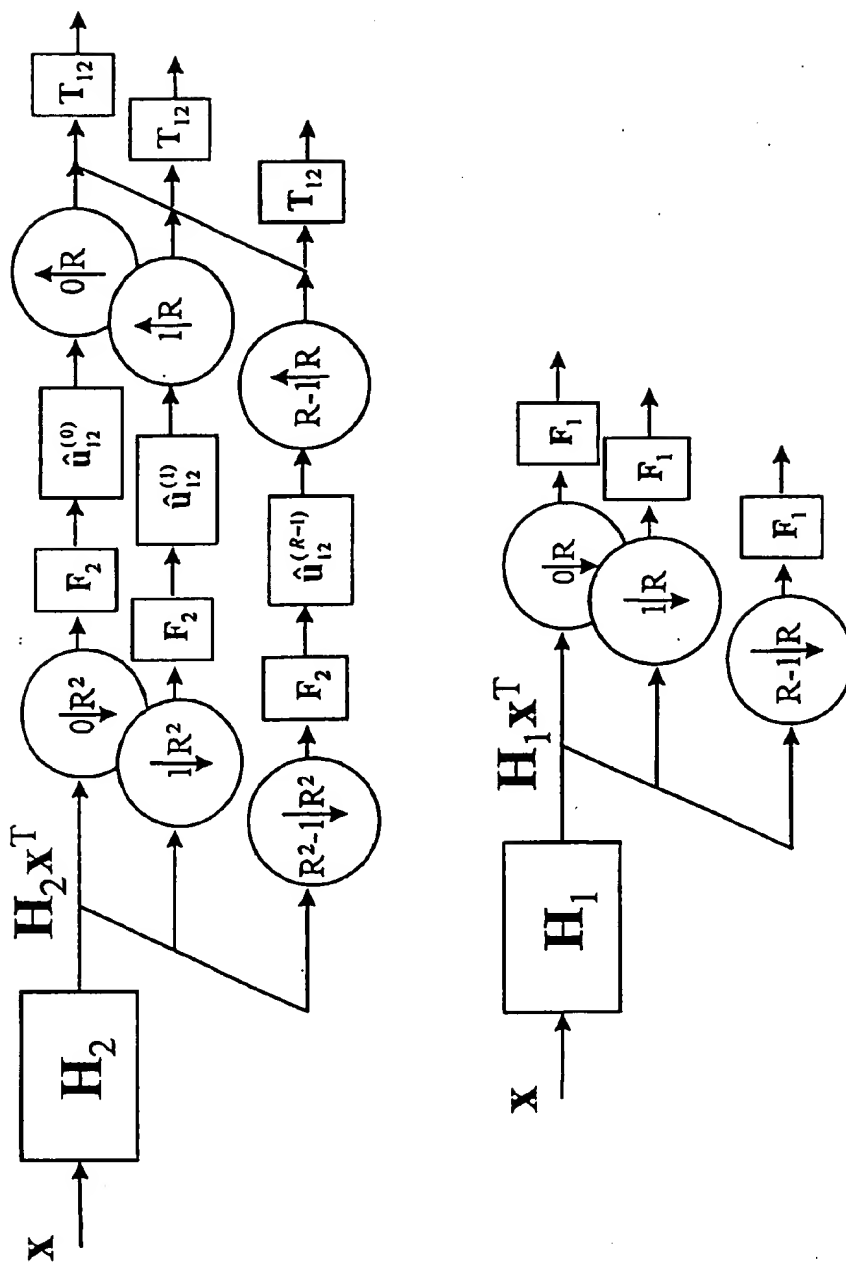


Figure 6

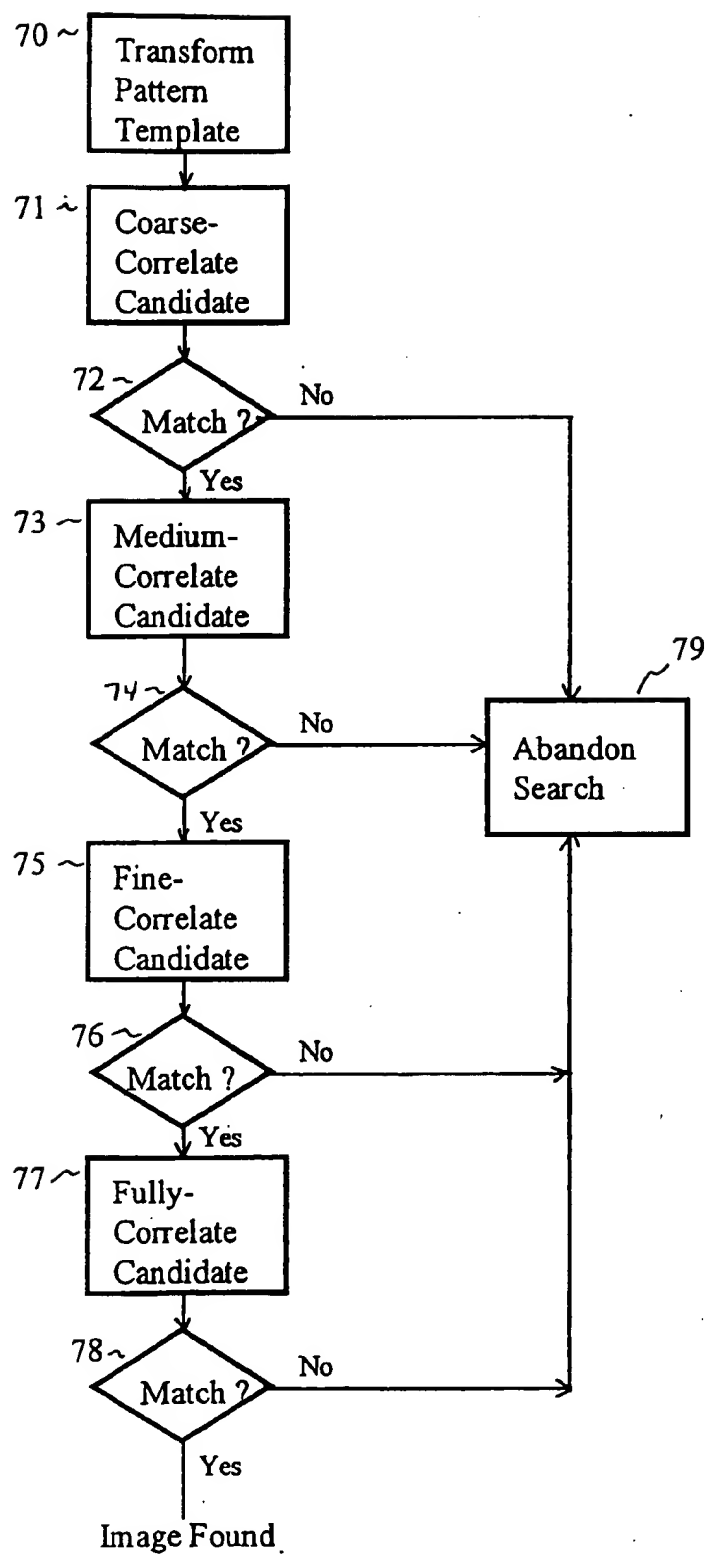


Figure 7.